

TITLE OF THE INVENTION

DATA COMMUNICATION CONTROL APPARATUS AND METHOD, AND  
DATA COMMUNICATION SYSTEM

5

BACKGROUND OF THE INVENTION

This invention relates to a data communication control apparatus and method for implementing a multipoint videoconferencing system by controlling the communication of video and audio signals among a plurality of terminals. The invention relates also to a data communication system that includes this data communication control apparatus.

Multipoint videoconferencing systems that are currently proliferating make it possible for videoconferencing terminals in compliance with the standard of ITU-T Recommendation H.320 to communicate with one another from multiple locations.

ITU-T Recommendation H.231, which defines the functional framework of multipoint connection devices and the form of multipoint communications, and ITU-T Recommendation H.243, which establishes an in-channel communications procedure, have been set up as standards relating to multipoint connection equipment for controlling the connections at the multiple locations of

videoconferencing terminals.

Fig. 6 is a block diagram illustrating the configuration of a multipoint videoconferencing system according to the prior art. The system includes a  
5 multipoint connection apparatus 122 compliant with H.231 and H.234, and videoconferencing terminals 121 compliant with H.320. The multipoint connection apparatus 122 interconnects the videoconferencing terminals 121 located at three or more points and supervises audio  
10 mixing, distribution of video data or multiple-screen synthesis of video data and chairperson control for facilitating conferencing. Thus, if the videoconferencing terminals are in compliance with H.320, a multipoint video conference connecting a  
15 plurality of remote locations can be realized by connecting the terminals via the multipoint connection apparatus.

However, an H.320-compliant videoconferencing terminal is very expensive because it is a dedicated  
20 terminal that necessitates high-load signal processing such as the compression and decompression of video data (ITU-T Recommendations H.261, H.263, etc.). Accordingly, it is difficult to realize a low-cost, low-power-consumption terminal that is capable of  
25 participating in videoconferencing, and to make such a terminal portable.

As a consequence of the foregoing, only a dedicated terminal can participate in conventional multipoint videoconferencing. In other words, a user can participate in conferencing only from a location at  
5 which a dedicated terminal has been installed.

In order to solve this problem, it has been proposed to provide the conventional multipoint connection apparatus not only with the function for interconnecting H.320-compliant videoconferencing  
10 terminals but also with a so-called voice gateway function that makes possible participation in a multipoint videoconferencing system by ordinary telephone.

Fig. 7 is a block diagram showing the configuration  
15 of a multipoint videoconferencing equipped with a telephone-based voice gateway function. As shown in Fig. 7, a plurality of dedicated videoconferencing terminals 131 and an ordinary telephone 133 are connected by a multipoint connection apparatus 132. In  
20 this system it is possible for the ordinary telephone 133 to participate in a multipoint video conference by voice only.

In a multipoint videoconferencing system of the kind illustrated in Fig. 7, it is possible to  
25 participate in a conference from an ordinary telephone but a sufficiently realistic sense of conference

participation is not obtained merely by telephone-based voice. In addition, since participation in a conference from a telephone is limited to audio, naturally it is impossible to participate in data conferencing in  
5 accordance with ITU-T Recommendation T.120.

#### SUMMARY OF THE INVENTION

Accordingly, it is an object of the present  
10 invention to provide a data communication control apparatus and method, as well as a data communication system, in which participation in a multipoint video conference by voice is possible from a general-purpose terminal.

15 According to the present invention, the foregoing object is attained by providing a data communication control apparatus for controlling data communication among a plurality of connected communication terminals, comprising: connecting means for connecting a general-  
20 purpose terminal; image generating means for generating image data that conforms to the general-purpose terminal; image distributing means for distributing the image data, which has been generated by the image generating means, to the general-purpose terminal via  
25 the connecting means; audio converting means for converting format of audio data in order that the audio

data may be communicated mutually between the general-purpose terminal and the plurality of communication terminals; and audio distributing means for distributing the audio data, whose format has been converted by the  
5 audio converting means, to the communication terminals and/or general-purpose terminal.

As a result, it is possible to view images and to communicate by voice with a communications terminal at a general-purpose terminal. Participation in a video  
10 conference from a general-purpose terminal thus becomes feasible.

Another object of the present invention is to provide a data communication control apparatus and method, as well as a data communication system, in which  
15 participation in a multipoint video conference by text chat is possible from a general-purpose terminal.

According to the present invention, the foregoing object is attained by providing a data communication control apparatus for controlling data communication  
20 among a plurality of connected communication terminals, comprising: connecting means for connecting a general-purpose terminal; image generating means for generating image data that conforms to the general-purpose terminal; image distributing means for distributing the  
25 image data, which has been generated by the image generating means, to the general-purpose terminal via

the connecting means; voice recognition means for recognizing voice data that has entered from the communication terminals and generating text data based upon this recognition; and data distributing means for  
5 distributing the text data to the general-purpose terminal.

As a result, it is possible to view images and to engage in a text chat with a communications terminal at a general-purpose terminal. Participation in a video  
10 conference from a general-purpose terminal thus becomes feasible.

Another object of the present invention is to provide a data communication control apparatus and method, as well as a data communication system, in which  
15 participation in a conference by text chat is possible from a general-purpose terminal even in a case where a dedicated communication terminal in a multipoint video conference is not equipped with a data conferencing function.

According to the present invention, the foregoing object is attained by providing a data communication control apparatus for controlling data communication among a plurality of connected communication terminals, comprising: connecting means for connecting a general-  
20 purpose terminal; image generating means for generating image data that conforms to the general-purpose  
25

terminal; image distributing means for distributing the image data, which has been generated by the image generating means, to the general-purpose terminal via the connecting means; voice recognition means for  
5 recognizing first voice data that has entered from the communication terminals and generating text data based upon this recognition; data distributing means for distributing the text data to the general-purpose terminal; voice synthesizing means for synthesizing  
10 second voice data based upon text data that has entered from the general-purpose terminal; and audio distributing means for distributing the second voice data to the communication terminals.

As a result, text data that as been entered from a  
15 general-purpose terminal is distributed to communications terminals as synthesized voice. Participation in a video conference from a general-purpose terminal thus becomes feasible.

The invention is particularly advantageous in that  
20 a general-purpose terminal can participate in a multipoint videoconferencing system in a simple manner.

Other features and advantages of the present invention will be apparent from the following description taken in conjunction with the accompanying  
25 drawings, in which like reference characters designate the same or similar parts throughout the figures

thereof.

#### BRIEF DESCRIPTION OF THE DRAWINGS

5       The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate embodiments of the invention and, together with the description, serve to explain the principles of the invention

10       Fig. 1 is a block diagram illustrating a multipoint videoconferencing system according to a first embodiment of the present invention;

15       Fig. 2 is a block diagram showing the construction of a multipoint connection apparatus in the system of Fig. 1 according to this embodiment;

      Fig. 3 is a diagram showing an example of the structure of HTML data in this embodiment;

20       Fig. 4 is a block diagram showing the construction of a multipoint connection apparatus according to a second embodiment of the present invention;

      Fig. 5 is a block diagram showing the construction of a multipoint connection apparatus according to a third embodiment of the present invention;

25       Fig. 6 is a block diagram illustrating a multipoint videoconferencing system according to the prior art; and

      Fig. 7 is a block diagram showing the configuration



of a multipoint videoconferencing system having a telephone-based gateway function according to the prior art.

5 DESCRIPTION OF THE PREFERRED EMBODIMENTS

Preferred embodiments of the present invention will now be described in detail in accordance with the accompanying drawings.

10 <First Embodiment>

Fig. 1 is a block diagram illustrating the basic configuration of a multipoint videoconferencing system according to a first embodiment of the present invention. As shown in Fig. 1, the system includes  
15 dedicated videoconferencing terminals 41a to 41d, which are terminals in compliance with the standards relating to videoconferencing stipulated by ITU-T Recommendations H.320, H.323 and H.324, etc., a videoconferencing multipoint connection apparatus 42 which, in addition to  
20 the functions relating to a conventional multipoint connection apparatus, also has a function for automatically generating HTML (HyperText Markup Language) data and an HTTP (HyperText Transfer Protocol) server function.

25 The system further includes general-purpose communication terminals 43a, 43b. Data such as video

and audio that has been stored in the HTTP server within the multipoint connection apparatus 42 can be displayed or converted to sound by the general-purpose communication terminals 43a, 43b. Personal computers or network computers incorporating a so-called WWW browser are applicable as the general-purpose communication terminals 43a, 43b. Recent progress in semiconductor technologies also makes it possible to use simpler transportable terminals.

10 Lines 44a to 44f connect the terminals to the multipoint connection apparatus 42. A public network such as an ISDN or PSTN and a LAN (Local Area Network) such as the Ethernet are applicable. It is assumed here that the general-purpose communication terminals 43a, 43b are connected to the multipoint connection apparatus 15 42 via the lines 44e, 44f, respectively, in accordance with the IP (Internet Protocol).

Thus, in this embodiment, the already existing videoconferencing terminals 41a to 41d and the general-purpose communication terminals 43a, 43b such as 20 personal computers or network computers are interconnected by simple means.

Fig. 2 is a block diagram showing the details of the multipoint connection apparatus 42 according to this 25 embodiment. As shown in Fig. 2, the apparatus includes network interface units 11a to 11c through which the

videoconferencing terminals 41a to 41d are accommodated in this system. These units supervise line interfaces of an ISDN or PSTN, etc. Multiplexer/demultiplexers 12a to 12c apply multiplex/demultiplex processing to various data relating to images, voice and control, etc. The multiplexer/demultiplexers 12a to 12c are in compliance with H.221, by way of example. The number of network interfaces and multiplexer/demultiplexers in the multipoint connection apparatus 42 correspond to the number of dedicated videoconferencing terminals capable of being connected to the multipoint connection apparatus 42.

Network interface units 20a, 20b are for connecting general-purpose terminals. As in the manner of the network interface units 11a to 11c for the dedicated videoconferencing terminals, these network interface units supervise interfaces of a public network, such as an ISDN or PSTN, etc., or interfaces of a LAN such as the Ethernet.

An audio processor 13 applies decoding/mixing processing to audio code data from the dedicated videoconferencing terminals 41a to 41d, thenceforth encodes the data again and distributes it to each of the dedicated videoconferencing terminals 41a to 41d and to a voice-communication protocol converter 21. An image switching unit 14 selectively switches among image

signals, which are distributed to the dedicated  
videoconferencing terminals 41a to 41d, in accordance  
with a control signal issued by a chairperson terminal,  
by way of example. A data distribution unit 15 executes  
5 processing for distributing T.120-compliant data  
conferencing data, etc. A system controller 16 carries  
out overall control of the system. An image converter  
17 generates still pictures by decompressing moving-  
picture code data that has been demultiplexed by the  
10 multiplexer/demultiplexers 12a to 12c.

An HTML data generator 18 converts, to HTML files,  
still pictures generated by the image converter 17 and  
various information relating to videoconferencing in  
this system. An HTTP server 19 provides various  
15 videoconferencing information generated by the HTML data  
generator 18 to the general-purpose communication  
terminals 43a, 43b connected to the multipoint  
connection apparatus 42.

The voice-communication protocol converter 21  
20 performs mutual conversion of encoding methods and  
protocols in order to realize mutual communication of  
audio data among the videoconferencing terminals 41a -  
41d and general-purpose communication terminals 43a, 43b  
connected to the multipoint connection apparatus 42.

25 A controller 30 has a CPU 31, a ROM 32 and a RAM  
33. The CPU 31 performs overall control of the above-

mentioned components in accordance with a control program stored in the ROM 32. The RAM 33 is used as the working area of the CPU 31.

The operation of the multipoint connection apparatus 42 in the videoconferencing system of this embodiment will be described in detail with reference to Fig. 2.

Multimedia multiplexed data that has been sent to the multipoint connection apparatus 42 via the dedicated network interfaces 11a to 11c accommodating the videoconferencing terminals 41a to 41d is demultiplexed to audio code data, video code data and low-speed data for data conferencing by the multiplexer/demultiplexers 12a to 12c. The demultiplexed audio code data is decoded by the audio processor 13 and mixed with data obtained by decoding other audio code data that has been demultiplexed by any of the multiplexer/demultiplexers 12a to 12c.

It should be noted that audio code data enters the multipoint connection apparatus 42 also from the general-purpose communication terminals 43a, 43b via the network interfaces 20a, 20b. This audio code data is subjected to a protocol conversion by the voice-communication protocol converter 21 so as to conform to the videoconferencing terminals 41a to 41d. The audio processor 13 decodes the protocol-converted audio code

data from the general-purpose terminals 43a, 43b, mixes this data with decoded data obtained by decoding the audio code data from the videoconferencing terminals 41a to 41d and then encodes the results.

5       The mixed audio code data is multiplexed with image data in the multiplexer/demultiplexers 12a - 12c and the multiplexed data is distributed to the dedicated videoconferencing terminals 41a - 41d. At the same time, the mixed audio code data is converted to a  
10       prescribed protocol by the voice-communication protocol converter 21, after which the resulting data is sent to the general-purpose terminals 43a, 43b via the network interfaces 20a, 20b.

15       Mutual communication of audio data is performed in real-time between the general-purpose communication terminals 43a, 43b and the multipoint connection apparatus 42 by RTP (Real-Time Transfer Protocol) of the IP. That is, the voice-communication protocol converter 21 performs a mutual conversion between the Internet  
20       protocol and the videoconferencing protocol in relation to the audio data.

      Among the various items of video code data demultiplexed by the multiplexer/demultiplexers 12a to 12c, the video code data from a terminal designated by  
25       the image switching unit 14 is sent to the multiplexer/demultiplexers 12a to 12c, whence the data

is distributed to the videoconferencing terminals 41a to 41d. The video code data demultiplexed by the multiplexer/demultiplexers 12a to 12c is sent also to the image converter 17 at the same time. The image  
5 converter 17 decodes the coded video data and converts one frame of video data to a still picture. The still picture obtained by the conversion is compressed by an encoding method stipulated by the JPEG (Joint Photographic Experts Group), by way of example.

10 The HTML data generator 18 generates HTML data based upon the still picture from each terminal generated by the image converter 17 and information relating to videoconferencing. The HTTP server 19 provides the HTML data generated by the HTML data  
15 generator 18 to the general-purpose communication terminals 43a, 43b, which possess an HTML viewer function, connected to the network interfaces 20a, 20b.

Fig. 3 illustrates the manner in which HTML files generated by the HTML data generator 18 are perused by  
20 an HTML viewer. As shown in Fig. 3, a window 51 indicates information relating to a videoconference. By way of example, the window 51 presents conference starting time and information relating to participants. Windows 52a to 52e display still pictures generated by  
25 the image converter 17, e.g., still pictures of the participants taking part in the conference by each of

the terminals. A window 53 is for T.120-compliant data conferencing. This window displays text chat, shared applications, etc.

Thus, in accordance with this embodiment, as  
5 described above, the user of a general-purpose terminal connected to a network interface via a network is enabled, by an IP-related protocol such as RTP, to perform voice communication with other dedicated  
10 videoconferencing terminals and is also enabled to peruse video from a dedicated videoconferencing terminal as a still picture by utilizing an HTML viewer.

Further, an ordinary personal computer or network computer having an audio input/output function can be employed as the general-purpose terminal that implements  
15 the above functions. As a result, a person can participate in a conventional multipoint video conference from any location via an ordinary telephone line and it also becomes readily feasible to participate in a conference from a transportable terminal.

20 <Second Embodiment>

A second embodiment according to the present invention will now be described.

Since the basic configuration of the videoconferencing system according to the second  
25 embodiment is similar to that shown in Fig. 1, which illustrates the first embodiment, it need not be



described again here. The multipoint connection apparatus 42 according to the second embodiment, besides having the functions relating to the conventional multipoint connection apparatus, is provided not only  
5 with the function for automatically generating HTML data and the HTTP server function but also with a function through which the content of a conversation with a person speaking at a conference is converted to text by a voice recognition mechanism.

10 Fig. 4 is block diagram showing the details of the multipoint connection apparatus 42 according to the second embodiment. Components in Fig. 4 identical with those of the first embodiment illustrated in Fig. 2 are designated by like reference characters and are not  
15 described again.

In Fig. 4, the audio processor 13 applies decoding/mixing processing to audio code data from the dedicated videoconferencing terminals 41a to 41d, thenceforth encodes the data again and distributes it to  
20 each of the dedicated videoconferencing terminals 41a to 41d and to a voice recognition unit 22. The latter subjects the audio data, which has been mixed by the audio processor 13, to recognition processing. The operation of the videoconferencing system according to  
25 the second embodiment will be described in detail with reference to the system configuration shown in Fig. 4.

Operational aspects similar to those of the first embodiment need not be described in detail again.

Multiplexed multimedia data that has been sent to the multipoint connection apparatus 42 via the network  
5 interfaces 11a to 11c accommodating the dedicated videoconferencing terminals 41a to 41d is demultiplexed to audio code data, video code data and low-speed data for data conferencing by the multiplexer/demultiplexers 12a to 12c.

10 The demultiplexed audio code data is decoded by the audio processor 13 and mixed with data obtained by decoding other audio code data that has been demultiplexed by any of the multiplexer/demultiplexers 12a to 12c. The mixed audio code data is coded further,  
15 after which the resulting data is multiplexed with image data in the multiplexer/demultiplexers 12a to 12c and then distributed to the dedicated videoconferencing terminals 41a to 41d. At the same time, the audio data that has been mixed by the audio processor 13 is sent to  
20 the voice recognition unit 22.

The voice recognition unit 22 subjects the input audio data to recognition processing and converts the obtained results to text data. The text data generated is distributed to the general-purpose terminals 43a, 43b  
25 as text-chat data in a data conferencing function in compliance with T.120, by way of example. It should be

noted that the data conferencing function mentioned here is one through which mutual communication of text-chat data is performed between the multipoint connection apparatus 42 and the general-purpose terminals 43a, 43b  
5 connected to the network interfaces 20a, 20b. Accordingly, text data that has entered from the general-purpose terminals 43a, 43b also is displayed in the text-chat window of the dedicated videoconferencing terminals 41a to 41d. The data distribution unit 15  
10 supervises distribution processing relating to this data communication. Meanwhile, in a manner similar to that of the first embodiment, video code data that has been demultiplexed by the multiplexer/demultiplexers 12a to 12c is decoded by the image converter 17 and converted  
15 to HTML data by the HTML data generator 18 so that it can be viewed in the manner shown in Fig. 3. In Fig. 3, the window 53 for data conferencing compliant with T.120, etc., displays, in the form of text chat, text data obtained in the voice recognition unit 22 based  
20 upon voice data that has entered from the dedicated videoconferencing terminals 41a to 41d. An example of the text chat is the content of remarks made by participants at a conference.

The users of the general-purpose terminals 43a, 43b  
25 are capable of entering their own remarks at a conference in the form of text displayed in the window

53. Entered text data is distributed to the dedicated  
videoconferencing terminals 41a to 41d. More  
specifically, the users of the general-purpose terminals  
43a, 43b are capable of mutual text-based communication  
5 with the users of the dedicated videoconferencing  
terminals 41a to 41d by way of the text-chat function.

Thus, in accordance with this embodiment, as  
described above, the user of a general-purpose terminal  
connected to a network interface via a network is  
10 capable of performing text-data-based mutual  
communication with a dedicated videoconferencing  
terminal by utilizing a text-chat function.  
Furthermore, video from a dedicated videoconferencing  
terminal can be viewed as a still picture by utilizing  
15 an HTML viewer. In addition, a personal computer or  
network computer having a comparatively handy functional  
architecture that does not include a multimedia function  
can be used as the general-purpose terminal for  
implementing the above-mentioned functions. As a  
20 result, participation in a multipoint videoconference  
from any location and from a transportable terminal is  
made easier.

#### <Third Embodiment>

A third embodiment according to the present  
25 invention will now be described.

Since the basic configuration of the

videoconferencing system according to the third embodiment is similar to that shown in Fig. 1, which illustrates the first embodiment, it need not be described again here. The multipoint connection apparatus 42 according to the third embodiment has a voice synthesizing function in addition to the function for converting the content of remarks to text by the voice recognition mechanism described in the second embodiment.

Fig. 5 is block diagram showing the details of the multipoint connection apparatus 42 according to the third embodiment. Components in Fig. 5 identical with those of the second embodiment illustrated in Fig. 4 are designated by like reference characters and are not described again.

In Fig. 5, a voice synthesizer 23, the input to which is text data sent from the general-purpose terminals 43a, 43b connected to the network interfaces 20a, 20b, executes voice synthesizing processing for converting the content of the text data to voice data. The synthesized voice data is mixed by the audio processor 13 with voice data that has been sent from the dedicated videoconferencing terminals 41a to 41d and the resulting data is distributed to the appropriate dedicated videoconferencing terminals 41a to 41d again. As a result, text-chat data from conference participants

using the general-purpose terminals 43a, 43b can be heard as voice information by the conference participants using the dedicated videoconferencing terminals 41a to 41d.

5        Thus, in accordance with the third embodiment, as described above, even if a dedicated videoconferencing terminal does not possess a data conferencing function such as a text-chat function, mutual communication with a general-purpose terminal connected to the  
10 videoconferencing system can be realized by synthesized voice.

      In each of the embodiments described above, it is explained that still-picture generation in the image converter 17 is performed at an arbitrary timing.  
15 However, it is also possible to start the conversion processing in response to commands from the general-purpose terminals 43a, 43b, by way of example. In such case the users of the general-purpose terminals 43a, 43b select the image windows 52a to 52e updated on the HTML  
20 browser, and the image converter 17 starts the conversion of the image and re-generates the still-picture data in response to an image request signal that conforms to the selection made, whereby the window display is updated. By thus generating a still picture,  
25 which is displayed in a window, in accordance with a timing designated by a user, even the users that have

connected from the general-purpose terminals 43a, 43b can share the necessary images in the course of a conference.

Further, in each of the foregoing embodiments, 5 examples in which HTML and an HTTP server are used as the means of supplying image data to the general-purpose terminals 43a, 43b are described. However, the present invention is not limited to such an arrangement. Any method of supplying image data may be utilized so long 10 as it is one through which hyperlinked multimedia data can be supplied.

Further, in regard to the voice communication protocol for voice communication with the general-purpose terminals 43a, 43b, any protocol may be used so 15 long as it makes real-time communication possible.

#### <Other Embodiments>

The present invention can be applied to a system constituted by a plurality of devices (e.g., a host computer, interface, reader, printer, etc.) or to an 20 apparatus comprising a single device (e.g., a copier or facsimile machine, etc.).

Furthermore, it goes without saying that the object of the invention is attained also by supplying a storage medium storing the program codes of the software for 25 performing the functions of the foregoing embodiments to a system or an apparatus, reading the program codes with

a computer (e.g., a CPU or MPU) of the system or apparatus from the storage medium, and then executing the program codes.

5 In this case, the program codes read from the storage medium implement the novel functions of the invention, and the storage medium storing the program codes constitutes the invention.

Further, the storage medium, such as a floppy disk, hard disk, optical disk, magneto-optical disk, CD-ROM,  
10 CD-R, magnetic tape, non-volatile type memory card or ROM can be used to provide the program codes.

Furthermore, besides the case where the aforesaid functions according to the embodiments are implemented by executing the program codes read by a computer, it  
15 goes without saying that the present invention covers a case where an operating system or the like running on the computer performs a part of or the entire process in accordance with the designation of program codes and implements the functions according to the embodiments.

20 It goes without saying that the present invention further covers a case where, after the program codes read from the storage medium are written in a function extension board inserted into the computer or in a memory provided in a function extension unit connected  
25 to the computer, a CPU or the like contained in the function extension board or function extension unit



performs a part of or the entire process in accordance with the designation of program codes and implements the function of the above embodiment.

5 As many apparently widely different embodiments of the present invention can be made without departing from the spirit and scope thereof, it is to be understood that the invention is not limited to the specific embodiments thereof except as defined in the appended claims.